Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 1 / ISSUE 3 / DECEMBER 2006 / 42 - 57

This paper is Available online at
www.jtaer.com

P2P's Significance for eBusiness: Towards a Research Agenda

Roger Clarke[1]

[1] Xamax Consultancy Pty Ltd, Canberra, Australia
and Visiting Professor in eCommerce at Uni. of Hong Kong,
in Cyberspace Law and Policy at Uni. of N.S.W. and in Computer Science at A.N.U.
Roger.Clarke@xamax.com.au

Received 25 April 2006; received in revised form 15 August 2006; accepted 13 October 2006

Abstract

Applications running over peer-to-peer (P2P) networks have exploded since the late 1990s. Research is needed into many aspects of P2P. These include architecture, application functionality, and the categories of digital works that are shared using P2P facilities. A range of significant strategic, consumer and policy issues also arise, such as challenges to the operation of copyright, defamation and other censorship laws. Organisations affected by P2P are devising and deploying countermeasures, such as technological protections for digital works and attempts to identify devices and users. These too require study. This paper presents the landscape of research opportunities, indicating methods that might be appropriately applied to the various kinds of questions.

Key words: P2P overlay networks, P2P applications, file-sharing, strategic impacts, policy impacts

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 1 / ISSUE 3 / DECEMBER 2006 / 42 - 57

This paper is Available online at
www.jtaer.com

1 Introduction

Applications of Peer-to-peer (P2P) technology have exploded since the late 1990s. After an initial period of confusion, the technology itself is becoming better-understood and better-documented. Its characteristics are very different from the predominant approaches of the 1980s and 1990s, however, and it appears to have a great deal of significance for corporate strategy and for public policy. Yet there remains a considerable degree of uncertainty and ambiguity about its applications, and its impacts and implications are far from being adequately understood.

As with any new domain, the research that has been conducted during the first few years has been opportunistic, and lacks structure and cohesion. A considerable amount of research has been conducted within the computer science discipline (e.g. [42]). On the other hand, it has until recently attracted far less attention within business-related disciplines, and from an eBusiness perspective. Searches in the IS literature identified few papers of relevance. For example, among the 450 papers published in Commun. AIS to April 2006, only one article directly addressed P2P [47], and it was focused on the needs of CIOs, not researchers.

The purpose of this paper is to propose a framework within which research into P2P's eBusiness impacts and implications can be conceived, managed and implemented. In developing this framework, study was first undertaken to establish the nature of a 'research agenda'. Literature surveys were conducted in a range of areas relevant to P2P. Particular attention was paid to attempts to impose structure on the topic. Research in computer science is well-developed, but to date far less attention has been paid to P2P in the formal literatures in information systems and other management disciplines. The work was informed by research previously conducted by the author into business models for eCommerce, by the author's consultancy assignments and experience as an expert witness, and by assessment of material on the web-sites associated with a wide range of P2P services. Drafts of the resulting proposals were exposed to colleagues, seminars were presented at four universities in three countries, and feedback was incorporated into later versions of the paper.

The paper commences by providing brief reviews of P2P's origins, nature and usage. It then discusses the concept of a 'research agenda', and proposes a framework. Each of the segments of the framework is then examined, in order to identify the kinds of questions that need to be researched, and the kinds of research techniques that can be employed in order to work towards answers.

2 The Emergence of P2P

This preliminary section provides a brief background to P2P, sufficient to lay the foundation for the analysis that follows. It considers in succession the origins of P2P, its nature, and the scale of P2P traffic.

2.1 Predecessors

During the 1940s and 1950s, computers were large standalone devices, operating in a single (large) room. During the 1960s, as techniques were developed to enable communications among devices, an architecture arose that is commonly referred to as 'master-slave': a central, relatively powerful computing device controlled simple devices installed remotely at the end-points of communication lines which were referred to as 'terminals'.

During the 1970s and early 1980s, processing power and storage capabilities were added to the remote devices, and gradually a new architecture emerged. A client (software running on one device) requested a service from a server (software running on another device). Servers originally ran on powerful central devices (typically, 'mainframe' computers), whereas clients ran on many, remote, less powerful devices (typically, PCs). The nature of coordination and control had changed: the masters had become servers, and the slaves had become clients.

Since the late 1980s, client-server architectures have been the norm [4] [45]. Most of our familiar working-tools currently operate in this way, including the downloading of emails from a mailbox and the use of a browser to access web-pages. Indeed, client-server architectures have been so dominant that many people who became computer users during the period 1980 to 2000 simply assumed that this was how computers collaborate, and did not appreciate that alternative architectures exist and that more could be conceived.

Client-server architecture is mature, and many sophisticated variants exist, featuring server dispersal, outsourced service provision, and replication and synchronisation of directories, file-catalogues and databases. All retain the characteristic of centralised control and coordination [8].

But substantial changes have taken place. The capacity of remote devices has continually increased, and the aggregate 'power at the edge of the net' has grown to be larger than that of the population of old-style hosts. Meanwhile, the extent of public accessibility to the Internet has exploded since the early 1990s. Initially, the devices whose power was being utilised were conventional 'PC's and workstations. Increasingly, other devices have become significant, including handheld devices (of many and various kinds including personal digital assistants – PDAs, and

some mobile phones, digital cameras, and networked gaming consoles), home entertainment centres, and even appliances (in the sense of 'white goods' such as printers and refrigerators).

In response, client-server architecture has been embellished. It is being applied to exploit the opportunity presented by accessible, unused processor-cycles and spare storage-capacity. The best-known initiatives have been collaborative processing services to search large data collections for patterns. A leading project in 1999 was the Electronic Frontier Foundation's DES cracker project which tested of a set of possible cryptographic keys [17]. Since 1999, the '@home' projects have been using free capacity in this way, initially SETI@home [1], but there have been many successors such as folding@home since 2000 and fightaids@home since 2003.

But client-server has come under challenge. In order to harness 'the power at the edge of the net', people began to experiment with an alternative architectural form. Now peer-to-peer architectures link millions of devices, some within organisations, but many controlled by consumers.

2.2 The Nature of P2P

This paper adopts the conventional view that the emergence of P2P is appropriately dated to the late 1990s, although brief reference is made later in this section to the fact that P2P was actually the re-discovery and re-badging of an existing idea, rather than a new invention.

Authoritative examinations of P2P are to be found in [35], [19], [40], [25 pp. 58-59, 75-78 and 136-145], [31], [42], [14], and other references in [43] and Wikipedia. Design principles for P2P file-sharing systems are provided in section 5 of [29].

An important precursor was the proposed Eternity service [2], but public awareness that change was occurring is associated with the Napster service. Napster was launched in 1998 and closed by court order in 2002 because it was found to breach copyright, and was unable to adapt its technology such that only legal downloads were feasible [12]. Conditioned by media focus during that period, public discussions still tend to be about high-volume applications used for transferring music and increasingly video, much of which has been in breach of copyright.

Napster was a hybrid of P2P and client-server architecture, in that, while the repository of music-files was highly distributed, the directory was centralised on the company's own servers. Many more sophisticated successors emerged, which avoided this 'single point-of-failure'. The most prominent names have been, in approximate succession of their release, Gnutella (in two distinct versions), the FastTrack network and associated Kazaa package [27], [28], [29], eDonkey, BitTorrent [15], [39], and Skype for real-time two-way audio/telephone. There are many other applications. Catalogues of P2P applications are maintained by various organisations, e.g. by Wikipedia.

Although the diversity of products is such that definition and classification of P2P is fraught with risks, a statement of scope is essential. Based on an examination of the literature, the following are identified as critical characteristics of P2P architecture:

- **server-functions are performed by many devices** (perhaps hundreds, perhaps hundreds of thousands, depending on the nature of the application and the volume of use). Each device needs to have significant residual processing and storage capacity beyond that necessary to support the user's own needs; and its connections need to have sufficient capacity, and to be persistent over a sufficiently long period of time, to justify the overheads involved in establishing and dis-establishing a server
- **any device that performs server-functions is able to perform all server-functions**
- **any device that is acting as a client is able to find one or more devices that are performing server-functions**; and
- **servers assist clients to find more servers**

In order to deliver an effective service, further practical requirements exist, in particular:

- a physical network of appropriate topology, reach and scale
- protocols for communication among nodes. A number of mature and effective protocols exist (see, for example, the catalogue maintained by Wikipedia), more are in development, and a considerable amount of research is being conducted into their design features
- naming conventions for nodes. Devices need to be identified, e.g. using a socket identifier (i.e. IP-address and port-number), or some application-specific naming mechanism
- naming conventions for digital objects, and for instances of digital objects. Relevant digital objects stored on the nodes need to be identified. A common approach is to generate a fixed-length hash of the contents of the file, and use that as part of the identifier for each instance of the object [16]
- metadata. Descriptive information about the processing service or digital object may be provided or generated, e.g. originator, publisher, title, date and version
- discovery mechanisms. A means is needed to locate target processing services and digital objects. Many current schemes rely on queries to multiple or successive servers until one or more instances of the target are found. A number of P2P architectures are built around a distributed hash table (DHT) [54]

- software that implements the necessary client and server functions. Many packages exist, and more are in development. See, for example, the catalogue maintained by Wikipedia
- the acquisition of participating devices. Server-resources might be contributed by one or more sponsors, by individuals donating them, or as a condition of use of the service. It is intuitively appealing to utilise a dispersed sub-set of the better-endowed among the available devices, but various P2P schemes apply various selection criteria

Drawing on and consolidating these elements, the following working definition is used in this paper:

> peer-to-peer (P2P) is a network architecture in which nodes are relatively equal, in the sense that each node is in principle capable of performing each of the functions necessary to support the network, and in practice many nodes do perform many of the functions

P2P architectures are argued to be capable of offering a number of advantages over conventional master-slave and client-server alternatives. Central among them is the avoidance of bottlenecks and single-points-of-failure, resulting in much-reduced dependence on individual devices and sub-networks. This enables improved resilience, much-improved scalability, much-improved ability to service highly-peaked demand, and resistance to denial of service attacks.

Applications are particularly likely to benefit from P2P architecture where demand is high relative to the power of individual processor-clusters and sub-networks, whether for an extended period of time, or just brief periods; and/or demand is widespread rather than arising in close proximity to individual processors and sub-networks. The sense in which the term 'proximity' is used may be related to terrestrial geography or network-topology. A fuller analysis is in [44].

As definitions of P2P emerged, it became clear that a number of important precursors existed. These include the Domain Name System (DNS), Usenet (now netnews) and even the despatch of email, which date from between the early 1970s and early 1980s. Indeed, the very first of the Request For Comment series of documents that define specifications for Internet protocols [41] has been argued to signal that peer-to-peer has always been part of the fabric of the Internet and even that peer-to-peer is the Internet's natural architectural form. The focus of this paper is, however, on post-1997 P2P technology.

2.3 The Scale of P2P Traffic

The first surge of growth occurred in 1998-2000, as a result of the success of Napster in attracting catalogue-entries for large numbers of music-files on consumers' disks. As Napster came under legal attack, the growth shifted to a succession of significantly more sophisticated alternatives.

As early as September 2002, it was estimated that 31 million Americans had shared music from a P2P service [22]. In 2004, the volumes of data being transmitted over P2P overlay networks were estimated by [23] as being about 10% of all Internet traffic, compared with 50% for the Web and 3% for all email including spam. Monitoring is conducted by a small number of companies such as BigChampagne, Cachelogic and TeleGeography, and a service called WebSpins is incorporated into Nielsen ratings in the U.S.A. [22].

The nature of P2P is such that traffic is difficult to define and to measure (e.g. Gong 2005). Transactions that depend on a central server or an identifiable set of distributed servers are easier to monitor than the actual traffic among widely distributed peers. As a result, there is considerable contention about traffic measurement.

A review of BigChampagne materials by a team at the Organisation for Economic Cooperation and Development (OECD) suggested that in early 2004 the number of simultaneous users of all P2P networks was running at close to 10 million, with some 50 million search queries per day. Around that time, the most popular network, FastTrack, appears to have provided access to about 2 million files, totaling close to 15 terabytes of data [33]. FastTrack had sunk from a peak of 5.5 million users and 60% of the market in mid-2003, to 4 million and 40% in early 2004. This was attributed to publicity about lawsuits by major music industry corporations in the U.S.A., and possibly to the early phases of a substantial decline in the quality of those files as a result of an apparent 'pollution attack' whereby the music industry introduced degraded versions of popular tracks to the network.

By mid-2004, the early dominance of audio files had been broken, with just under 50% of files appearing to contain audio, 25% video, and 25% other formats including software. Both audio and video files are large, video files particularly so. As would be expected, there was a modest correlation between P2P usage and the penetration of broadband. (Except as noted, all of these figures are from [37], which drew on various sources).

The patterns have continued to change. Sampling data from mid-2005 indicated that P2P represented 50-65% of "downstream traffic" (i.e. from ISPs to endpoint devices) and 75-90% of "upstream traffic" [6]. The different proportions arise because, with hitherto dominant client-server applications (such as the Web), the volume of upstream traffic is much smaller than that of downstream traffic. This is significant in several ways, not least

because most broadband services (such as ADSL and particularly cable) offer much smaller upstream capacity than downstream, because they were designed in an era in which client-server architectures were dominant.

The proportions of traffic that P2P represents vary considerably, however, in at least two ways. Firstly, the Web still dominates on of the c.300 backbone providers worldwide [49]. And secondly, the patterns of usage are highly variable across regions, with eDonkey dominant in South Korea, BitTorrent in the rest of Asia, and both FastTrack and Gnutella much-used in the USA, Canada and Western Europe [6]. Video has been dominating the volume since 2004, although the file-count may still be dominated by audio. But many other file-types are also represented, including software and backup data-files.

3 The Concept of a Research Agenda

The previous section provided an outline of P2P, noted a considerable number of uncertainties, and implied the need for research. The purpose of this paper is to provide a structured review of the areas within which research into P2P's significance for eBusiness needs to be conducted.

The term 'research agenda' is much-used to refer to such a structure, but seldom clearly defined. Some guidance is provided, however, by [50], [52] and [42].

Drawing on those sources, it is argued that a research agenda should provide the following elements:

- comprehensive coverage of the research domain under consideration
- a small set of key factors organised so as to provide at least a framework, and to the extent possible a taxonomy
- an indication of research already undertaken in each of the resulting quadrants or categories, and of needs and opportunities; and
- a discussion of appropriate research techniques to address the identified opportunities

In the case of P2P, it is of particular importance to take into account the layered nature of the phenomenon, because it is very likely that research opportunities and appropriate research techniques will be significantly different at each layer. It is proposed that two networking layers be differentiated, an applications layer, and two layers addressing impacts and implications, as follows:

1. **the underlying communications infrastructure**. This commonly (but not necessarily) comprises TCP sessions and/or UDP transactions, over IP, transmitted over wired and 'unwired'/wireless connections
2. **the 'overlay network'**. This is the virtual network that is provided by the enabling elements of P2P software
3. **applications**. These use the overlay network for discovery, and for the transfer of digital objects, streaming, or collaborative processing
4. **impacts**. P2P has substantial impacts on the strategies of corporations and government agencies, and on consumers
5. **implications**. P2P has important second-order implications for industry sectors, the economy and society. It is consequently giving rise to a considerable amount of policy activity. Countermeasures adopted by affected parties add to the implications and to the policy needs

The remainder of the paper is organised in accordance with the above 5-layer structure, and addresses each of the elements of a research agenda that were identified above. Because it is essential that the whole of the research domain be addressed, it is inevitable that a research agenda of article length will treat each segment fairly superficially.

4 The Underlying Communications Infrastructure

The primary network (or, more correctly, 'inter-network') over which P2P is implemented is **the open, public Internet**. Research questions include:

- are the present specifications of the TCP, UDP and IP protocols appropriate for P2P? Could adaptations significantly improve the performance of P2P applications?
- to what extent does P2P fulfil its promise of scalability, and coping with surges in demand? Is it effective only where the demand is network-dispersed (e.g. the Olympics, or a news item of world-wide interest), or also where the demand is network-localised?
- what is the susceptibility of P2P to stressful circumstances, such as warfare, during and in the aftermath of natural disasters, and under denial of service attack? Differently phrased, what is the downward scalability of P2P?

The open, public Internet is only one possible infrastructure for P2P. Similar questions arise in relation to **intranets and extranets**, and similar research techniques can be applied to test, for example, the efficacy of P2P-based backup and recovery within a corporate network, and the minimum network-scale needed to be effective.

Although the Internet Protocol Suite has been the primary focus of most designers and of most observers, other intermediate-layer protocols may also be appropriate hosts for P2P applications. Tests could be performed on

networks running other protocol suites, including large regional networks, large dispersed corporate networks, and small local area networks (LANs) such as Ethernet segments within a single small building or floor.

Testing is needed of the robustness of networks (i.e. to what extent can they continue to perform their functions under adverse circumstances?), and of their resilience (i.e. how easily can they be recovered after a major outage?). The possibility exists that future adaptations of network infrastructure might include features intended to present barriers to the 're-booting' of P2P networks. Hence examination is also needed of the susceptibility of networks to a coordinated denial of service attack.

Several research techniques are applicable to the research questions that arise in this segment of the P2P domain. Existing systems can be studied, and the scope exists for field experimentation and quasi-experimental designs (e.g. by adapting an existing P2P application to gather additional data, and to perform somewhat differently from the intentions of its designers). New systems can be constructed in order to trigger features of the underlying infrastructure. Based on an understanding of Internet characteristics, simulation models can be constructed, in order to predict behaviour under various P2P-generated conditions. A considerable amount of research has been conducted into such matters within the computer science discipline (e.g. that catalogued by [43] and [42]), and eCommerce researchers need to make themselves aware of, and draw on, that research.

5 The P2P 'Overlay Network'

P2P schemes generally involve a layer of software that intermediates between the application and the underlying telecommunications infrastructure. The function of this layer is to manage the linkage between nodes. This brings into existence a virtual network of nodes that is commonly referred to as an 'overlay network'.

There are a number of such schemes. Some are interwoven with applications, and many references fail to distinguish between the two. Well-known overlay networks include the original Napster, the original Gnutella, FastTrack, Kademlia, and eDonkey's Overnet. Catalogues of overlay networking tools are available at the web-sites off Wikipedia, Internet2 and Slyck. It has been demonstrated that a P2P overlay network can be implemented in a very small program [18], [46].

5.1 Technical Features

A first set of research questions relates to whether the claimed technical advantages over client-server architecture are realised in theory, and in existing, observable schemes. Investigations are likely to lead to sub-questions, such as whether the advantages only emerge once a particular scale is reached, whether there are limits to the scalability even of P2P, and the extent to which performance is dependent on the characteristics of participating devices, detailed design decisions and operational parameters. A survey is provided in [3].

Further questions arise in relation to various technical challenges. One cluster relates to the means whereby nodes discover other participating nodes. Searches for content may be based on metadata describing the desired file or stream, or on hashes of the file-content. More advanced forms of searching are being devised, to utilise taxonomies in order to approximate semantics (e.g. [36]). Research questions include the effectiveness of various discovery techniques for various forms of content under the various operational conditions of P2P networks.

An especially important technical aspect is the robustness of P2P networks. They are known to be vulnerable to masquerade attacks (falsified digital objects and services that purport to be known ones) and pollution attacks (adapted versions of known digital objects or services). A P2P overlay network is also more vulnerable to attack if its topology is relatively stable; but the more dynamic the topology, the greater the overheads of network management, and the smaller the percentage of requests that are satisfied quickly. The trade-offs under different approaches need to be evaluated.

Techniques that are particularly suited to the examination of such questions include field experimentation and quasi-experimental designs, laboratory experimentation, simulation, and engineering construction and de-construction.

5.2 The Approach Taken to Coordination

It is feasible to treat P2P as though it were a single architecture. On the other hand, there are already some distinct variants, and their technical advantages and disadvantages may vary considerably. Of especial interest is the approach taken to coordination within P2P overlay networks. Generally, the business of the network (most commonly file-discovery and file-transfer, but also streaming and collaborative processing) is conducted peer-to-peer directly between participating nodes, in some cases 1-to-1, and in other cases many-servers-to-1-client. The management of directory entries, and the various network management functions, may also be conducted in the same openly collaborative manner, or they may involve some specialisation among nodes, and perhaps some degree of centralised control.

This results in the following taxonomy of schemes for P2P overlay networks:

- **'pure P2P'**. All functions, including not only the storage of all relevant digital objects, but also the management of the directory and of the network, are distributed across a great many nodes, such that no node is critical to the network's operation; and hence no node can exercise control over the network. Examples include USENET, Fidonet, Freenet and the original Gnutella
- **'partially centralised P2P'**. The directory may be subject to a degree of control as a result of being hierarchically structured, as in the Domain Name System (DNS). In the most extreme cases, control over the operations of the file-catalogue may be centralised, as in the cases of Napster and in a different manner BitTorrent. Napster no longer exists because its centralised catalogue represented a 'single point of failure', which was subjected to a form of 'denial of service' attack (although by lawyers for aggrieved copyright-holders, rather than by 'hackers'); and
- **'two-tier P2P'**. Between the two extremes, the catalogue and some management functions may be substantially distributed but not fully distributed. This might create some scope for influence by some party. Important examples include FastTrack and hence Gnutella2

In addition, at least three special circumstances exist, which might create the possibility of some degree of control being exercised by some party. These are:

- the 'bootstrapping' phase of the overlay network, by which is meant the means whereby it is brought into existence at the outset, and on any subsequent occasion after it has collapsed and needs to be re-initialised. It is difficult to conceive of a means whereby P2P implementations could completely avoid some degree of centralisation during this phase. On the other hand, it has been suggested that the FastTrack network might run for years after the software was withdrawn, whether or not attempts were made to close it down by copyright-owning corporations, government authorities, or even the original sponsor
- the release of a new digital object. The opportunity might exist for an interested party, such as a copyright-owner or a censor, to intercept the initial publication of an object, and prevent its replication; and
- intrinsic software-update functions. The software that manages a P2P overlay network might be designed with an automated update facility (in effect a form of backdoor, trapdoor or trojan horse), which might create scope for control either by the party that originated the package, or that now controls the software updates, or by some other party

Research questions arise in relation to the operational characteristics of each of the architectural variants, and their relative controllability, and vulnerability. That leads to questions as to what contextual factors determine which of the various architectural alternatives is most appropriate to a particular application.

5.3 The Approach Taken to Ensuring Participation

A further consideration that appears to be vital to the success of P2P networks is the means whereby sufficient resources are acquired from participants. Most implementations of P2P architectures include features to encourage the availability of devices to perform as servers, and to discourage devices from 'free-riding' or 'over-grazing' the 'commons', which is a risk intrinsic to the architecture. Background on the economics of P2P networks is in [9].

One fairly common design feature is to ensure that nodes are not called upon to contribute resources for long periods of time, because that way the user is less likely to adopt measures to cause the server component to cease operating. This is a key technical reason why many P2P schemes involve ephemeral servers, a highly dynamic network topology, and hence highly volatile metadata. For the same reason, means are usually included for ensuring that nodes remain responsive to processes initiated by the device's local user(s).

Another rational concern among users is that the software that provides them with access to a P2P network may include functions that are harmful to the individual's interests. This is further discussed under Consumer Impacts below. Incentives to make resources available must be sufficient to overcome such 'distrust' factors.

5.4 Developer Philosophy

An interesting area of overlap between P2P and other research domains is the philosophy adopted by the developers of protocols, of standards, and of software. Some overlay network protocols and libraries are closed and proprietary (e.g. FastTrack, Altnet/Joltid and Skype), whereas others are open specifications (e.g. DNS, OpenNAP, BitTorrent and OpenFT).

Research questions include: Are there advantages for one or the other approach in terms of such features as speed of implementation, speed of adoption and quality assurance? Such questions might be initially addressed through conceptual research, simulation modelling and scenario-building. Scope also exists for field studies, but control over confounding variables is likely to prove very challenging.

Journal of Theoretical and Applied Electronic Commerce Research
ISSN 0718–1876 Electronic Version
VOL 1 / ISSUE 3 / DECEMBER 2006 / 42 - 57

This paper is Available online at
www.jtaer.com

6 P2P Applications

Built over the underlying communications infrastructure and the overlay network are applications of direct interest to users. One classification scheme for applications is provided by [47] p. 101. The research questions identified below address variously the categories of services that P2P can provide, and the features the software offers in order to deliver value.

6.1 Application Categories

Four broad categories of application can be identified, based on the nature of the resources that are being shared:

- **collaborative processing**. Possibilities include large-scale, brute-force numerical methods (e.g. in meteorology, and fluid dynamics experiments), and multi-player networked gaming. A catalogue of initiatives in these areas is at O'Reilly OpenP2P (Distributed Computation)
- **streaming**. This enables near-immediate access at the receiving end, and is therefore important for very large files, and transmissions of live events (e.g. [53], [21]). There are some successful applications, including some counter-intuitive ones such as PPLive for TV transmission in the People's Republic of China. Because of the limited extent to which IP multicasting has been deployed, peercasting has emerged as a potentially valuable category of application, with products including FreeCast, Octoshape, PeerCast, Alluvium and IceShare
- **message transfer**. This may be variously for cooperative publishing, for text-based conferencing/chat/instant messaging, and for audio/telephonic and perhaps video interactions. A catalogue of initiatives in these areas is at O'Reilly OpenP2P (Messaging Frameworks); and
- **file-sharing**. This has attracted most of the attention to date, among both developers and observers. Because of its widespread usage, and its clear relevance to eBusiness, research in relation to file-sharing is the primary focus of the remainder of this paper

P2P applications involve many different **categories of digital object**. These have rather different characteristics, and hence give rise to rather different impacts and implications. They include the following:

- **entertainment materials**. These are variously in text, image, sound and video formats. This appears to be the primary interest of users of the most popular packages and services
- **games data**. Examples include scenes, battle configurations, characters/avatars, and 'cheats';
- **learning materials**. These are variously in text, image, sound and video formats. This is the primary focus of Edutella and Lionshare
- **announcements**, of many different kinds. These include technical information, business information, entertainment, sports results, promotional messages and advertisements
- **news reports**. These are not limited to those generated by news organisations. They also include informal traffic emanating from members of the public at the scene of an event such as a concert, product-launch, sports tournament, demonstration, or crime-scene. Freenet targets the distribution of informal news reports [10]. A related topic is the wearing of video-capture and transmission devices by event participants [34];
- **message archival**. Beyond netnews, this may be for conferencing/chat/instant messaging, or for cooperative publishing. This is a target of Groove
- **emergency-services traffic**. P2P may have an important role to play in supplementing amateur radio during and in the aftermath of natural disasters such as earthquakes and tsunami. The necessary characteristics of such applications are likely to have a close relationship with those for operational military traffic
- **data backup and recovery**. The feasibility of this application is discussed in [44]
- **software releases**. Access to these may be highly-peaked for functional reasons, or because of popularity or fashion
- **software fixes/patches, and data such as virus signatures**. Digital objects associated with security vulnerabilities are a potentially highly valuable application of P2P. This is because they need to be accessed particularly quickly by very large numbers of devices, and hence generate highly peaked activity

The majority of P2P traffic during the period 1998-2003 appears to have been audio files, with video file sharing increasing as network and tail-end capacity has increased. Research is needed into usage patterns. Of particular interest is the provision of estimates of the proportion of file-sharing that is copyright-infringing, and the extent to which the infringements are for commercial and for consumption purposes.

The classification schemes for the nature of the shared resources are open-ended, and new services may be enabled by P2P networks. For example, P2P could be applied to support the reticulation of knowledge within organisations and beyond them, and to build recommender systems for collaboration among consumers.

Further research questions of considerable interest include: Is each P2P technology equally applicable to different categories of application (file-sharing, streaming, and collaborative processing) and different categories of digital object? Or are fundamentally different designs and hence separate architectures necessary? Will P2P, streaming services and Grid Computing develop separately, cross-fertilise, or merge?

Such questions can be addressed using non-empirical techniques such as conceptual research, simulation and scenario-building. Field and laboratory experimentation using existing applications and networks are also feasible, as are engineering construction and de-construction techniques.

6.2 Application Features

File-sharing applications are attracting more attention at present than message-transfer, streaming and collaborative processing, and are more mature. This section accordingly focuses on the features of packages intended for file-sharing.

Some packages are designed to utilise one specific overlay network; but many implement multiple protocols in order to provide access to several overlay networks. Researchers need to appreciate this distinction, and take it into account in their research designs, and in their interpretation of their sources and their data.

Multi-network packages are tending to dominate, and include Kazaa Media Desktop (KMD), Grokster, Morpheus and iMesh. Examples of specific-network packages include BitTorrent, PeerEnabler, MLDonkey and eMule. Catalogues of currently available packages are available at Wikipedia, Internet2, Slyck, O'Reilly OpenP2P (Distributed Search Engines) and O'Reilly OpenP2P (File-Sharing).

With the proliferation in applications, information is needed about patterns of usage of the applications. Research questions include: Which packages are used by what categories of user? Which packages are used for what categories of file? What forms of specialisation are offered? Which specialisations appear to be valued by which users, for which categories of file? These are empirical questions, which may be answered through field studies, supplemented by surveys and perhaps focus groups, and secondary research utilising postings to mailing-lists and bulletin-boards.

Deeper information is needed about application functionality. There are some basic functions that all applications need to perform, such as the provision of interfaces with the user, with the relevant overlay network(s), and with tools to render the files that are received (such as a web-browser and Adobe Acrobat for text and image content, and Windows Media Player and Apple QuickTime for audio and video). Some may themselves perform rendering functions. Many also offer management facilities for the files that are downloaded, including indexing, annotation and metadata management. Most also include the capacity to perform server functions. All need to perform operational functions such as maintaining a list of nodes in the overlay network that they can make contact with. At least some are likely to include administrative functions, such as the collection and reporting of statistics.

The functionality of each P2P application, and in some cases of each overlay network layer, could be expected to reflect the kinds of uses and users that the designers had in mind. An example is censorship-resistant publishing. To achieve all of the objectives, multiple tools might need to be used in conjunction with one another, such as Freenet [11], Publius [51] or some other anonymous authoring and publishing scheme, and Mixmaster [7] or some other anonymous-proxy method for posting the document onto servers. To date, there tends to be a trade-off between robustness and nymity, on the one hand, and user-friendliness, on the other.

Research questions include: What alternative approaches are adopted to application design (such as the use of existing tools versus custom-building)? What tools are popular among designers (e.g. are Internet Explorer and Windows Media Player commonly required, or is scepticism about Microsoft products apparent among user communities)? Do packages appear to be specialised to particular markets, or to be generalised to support wide varieties of users and file-types? Do packages incorporate means whereby data can be gathered about users, or uses? Do packages incorporate means whereby some party may be able to exercise control over aspects of the network's operation?

Engineering de-construction techniques are especially applicable to such questions. Field and laboratory experimentation may be valuable in conducting comparisons among applications. Interviews with developers might be a valuable supplementary technique.

Beyond the technical issues, the technology-in-use needs to be studied, to answer questions such as: Do users find the trade-offs selected by the designers to be appropriate? Such questions are appropriately addressed using surveys, and some of the interpretivist techniques.

7 Impacts

P2P schemes appear capable of having substantial impacts both on organisations and individuals.

7.1 Strategic Impacts

A variety of organisations may be affected, including:

- Internet Access Providers, as traffic flows become more symmetric, and users abandon contemporary asymmetric connection mechanisms via 56 kbps modem, ADSL (Asynchronous Digital Subscriber Line) and cable in favour of SDSL (Synchronous Digital Subscriber Line) and similar
- technology services providers, if, for example, demand for large centralised hosts declines, as more services migrate to dispersed devices
- organisations that are susceptible to the dissemination of confidential information, including not only trade secrets but also evidence of criminal behaviour such as corruption and collusion, and historical revisionism; and
- government agencies responsible for censorship, variously on the grounds of criticism of monarchs; seditious political comment; sexual content; incitement of violence and hatred; instruction in violence; and trade in proscribed goods and services

To date, however, the primary category of organisation deeply embroiled in the turmoil appears to have been corporations whose revenues and profits are significantly dependent upon copyrighted materials, in particular music publishers, and now publishers of video-format materials such as feature-films and documentaries. They are confronted by a succession of challenges:

- unauthorised reticulation. Digital objects can circulate in breach of the rights of the persons that own copyright in them. These may be 'born digital' objects, or physical objects that have been digitised without authorisation;
- unauthorised adaptation. Digital objects may be substantially altered (e.g. in order to add value, or to discredit the originator), or subtly altered (e.g. in order to circumvent protections, or to make the object more difficult to find)
- the identification of copyright objects. This is particularly the case where countermeasures of various kinds are adopted
- tracking of the movement of objects
- the identification of devices that store those objects and that traffic in them
- the association of the organisation or person responsible for a breach with the device that was used to perform the act that constitutes the breach
- the location of the responsible organisation or person
- the bringing of a suit against that party (which can be challenging because of trans-jurisdictional limitations)
- the gathering and presentation of evidence sufficient to win even civil, let alone criminal cases; and
- where cases are won, the proposal of interventions that could reasonably be awarded by court injunction. P2P has far fewer chokepoints at which either technical or legal measures can be aimed in order to counter unauthorised acts and overcome the many difficulties outlined above. In effect, the removal of a device as a result of the execution of a warrant or injunction is indistinguishable from other forms of denial of service attack, and the network tends to 'route around it'. The scope for interventions to have harmful side-effects is also very high. These issues are at the heart of court-cases running in 2004-05, as the music industry attacks Kazaa and others (e.g. [24])

With the challenges to enforceability comes a reduction in user accountability, because people perceive themselves and their behaviour to be difficult to identify, track and respond to. Traceability and the consequential possibility of retribution are not the only factors involved in control over anti-social behaviour; but they are very likely to be important factors.

The questions that arise in this segment of the P2P research domain are very different from those discussed in previous sections. Some are concerned with **the interplay between the behaviour of users and the law**. What proportion, and what volume, of file-sharing is conducted without an appropriate license? What proportion, and what volume, of file-sharing are in breach of copyright law? (The two questions are not equivalent). What is the incidence of unauthorised adaptation of files? What proportion of downloads involve active endeavours to avoid identification (such as the use of proxy-servers)? These can be addressed using field study, secondary data research, and perhaps field experimentation and engineering techniques.

Other questions involve study of **human factors**. For example: What proportion of the population comprises inveterate anti-capitalists whose behaviour is independent of the law and of public information campaigns? To what extent is consumer behaviour changed by the knowledge that legal action is being taken by copyright-owners? Has consumer payment morality changed since 1998? What is the elasticity of demand for various kinds of digital objects? Are there threshold prices for various categories of digital objects, above which consumer behaviour changes significantly? How do consumers perceive the use of pseudonyms by others, and by themselves? These are capable of being studied using surveys and by most forms of interpretivist techniques.

A further, important set of questions requires understanding of **financial aspects of business strategy**: What margins did pre-P2P prices offer copyright-owners in various contexts? What scope do they have to reduce costs? In each sector affected by P2P, are there examples of re-engineered corporations that have achieved significantly lower cost-profiles than their competitors? If so, are there factors that prevent other corporations in those sectors following their lead?

A particular cluster of issues arises in relation to **business models**. Prior work on business models is documented in [13]. Some of the important questions in need of research include: Do P2P technologies and applications

undermine some conventional business models, or demand their adaptation? Do they create the possibility of new forms of business model? Do business models exist that involve revenue-generation or cost-offset approaches different from those used by mainstream copyright-owning corporations in the sector? In particular, is it feasible for publishers to extract sufficient revenue from commercial users (such as radio stations and video-streaming services) such that they can afford to forego control over personal copying and use?

Does P2P create significant new scope for **dis-inter-remediation**? In particular, can copyright-owning organisations themselves exploit P2P technologies, as many theorists, and the changes brought about by Napster, and more recently Apple's iTunes, Kazaa and BitTorrent, have encouraged them to do? Might publishing corporations themselves be dis-intermediated out of existence, with originators communicating directly with downloading services and streaming channels, and/or directly with consumers? If so, how long might such a transition take, and can large publishers adapt quickly enough to survive?

7.2 Impacts on the Consumer/Citizen

The primary impacts on consumers are positive. Files are much more readily and quickly accessible, and in most cases to date there is no direct cost involved. On the other hand, consumers do incur costs for their own infrastructure, network connection and possibly traffic; they pay indirectly to the extent that advertising costs force up the prices of other goods and services; and, after a very long gestation period, early movers like Apple, through iTunes, are finally forcing commercial operators to apply forms of P2P that incorporate direct charging models.

Security vulnerabilities exist until learning has taken place, appropriate features have been designed and implemented, and new versions of software have been widely deployed. One concern is that consumer devices may be subject to surreptitious enlistment into P2P networks, without the informed and freely-given consent of the person responsible for the device, or perhaps without any kind of consent, or even without the person's knowledge. Research questions include: What disclosures are made to users of P2P applications regarding the use of their devices as servers? To what extent are consumers aware that their P2P application may perform server as well as client functions? What protections exist to prevent P2P-server activity significantly affecting the consumer's use of the device?

Because of their popularity, P2P applications have also been attractive to people seeking vectors for the various kinds of 'malware' such as 'trap doors', 'backdoors' and 'trojan horses', including the acquisition of so-called 'zombies'. In addition, some P2P applications include adware (which uses data on or about the device or its presumed user to select among alternative advertisements to be displayed on the device), and some include spyware (which extracts personal data from the device and sends it elsewhere). It is also feasible that P2P applications may contain features to modify or delete data or software on the device. Research questions include: What evidence exists of the use of P2P applications as vectors for malware?

Negative impacts on individuals arise where P2P networks are used to reticulate digital objects against the wishes of people who have interests in them. These interests vary greatly. For example, in some cultures, pictures of deceased persons are sacrosanct; and in others veneration of the aged is important (even if those aged have a dubious past). The interests in secrecy that tend to dominate discussion are the protection of a person's reputation, the avoidance of civil lawsuits, and the avoidance of criminal prosecutions. These are far more than merely a personal and psychological matter. They have a social and even a political dimension, because they might increase conformity of behaviour with whatever the society perceives to be 'normal', might work against freedom of political speech, and might act as a disincentive against people being prepared to offer themselves for public office. Research questions include: Is there evidence of P2P being applied to character assassination, the embarrassment of hypocrites, and the exposure of villains?

8 Implications

P2P is also having second-order effects on industry sectors, and on governments. This section scans these broader issues, in order to identify opportunities for researchers to contribute to understanding of the P2P phenomenon.

8.1 Corporate and Industry Sector Implications

Organisations have interests in information flows being restricted, in such contexts as trade secrets, commercial negotiations, insider trading, and the avoidance of accusations in relation to environmental or social depravity, collusion and corruption. Research questions include: Is there evidence of P2P being applied to corporate leaks and whistleblowing?

The issue also arises of the dissemination of incomplete, misleading, and utterly false information, variously in order to harm reputation, drain an organisation's or person's resources, chill their behaviour, or inflate stock-prices. Research questions include: Is there evidence of P2P being applied to the spreading of rumours?

More broadly, those industry sectors appear to be under threat that are dependent for their revenue on their ability to control the dissemination of objects that are the digital or digitisable. Rapid change is harmful to investors, to employees, to organisations upstream and downstream from those sectors, and to regions dependent on them for employment. There are economic and social interests in industry sector re-structuring being gradual rather than sudden.

Research questions include: What is the spread of returns from copyright-dependent industry sectors to originators, publishers, publishers' contractors, and elements of the distribution chain? What changes in turnover, profitability and employee-count have been apparent within copyright-dependent industry sectors? To what extent are copyright-dependent sectors large employers in depressed economic regions?

8.2 Implications for Governments

Censorship laws may be undermined by the use of P2P networks. Research questions include: Are P2P networks being used for proscribed content, such as 'seditious' materials, child pornography, incitement to violence and hatred, and trading in proscribed objects such as scheduled drugs, explosives, firearms, Nazi memorabilia, and goods deriving from threatened species?

In varying degrees, and using varying methods, governments of all nation-states seek to control information and public perceptions. They may find their capacity to do so undermined by the use of P2P networks. Research questions include: Are P2P networks being used to achieve citizen-driven Freedom of Information to supplement existing, narrow government mechanisms? Is there evidence that government 'propagandists' and 'spin-doctors' are now having less success in guiding and manipulating public opinion?

Of particular interest is censorship of information about current events, or 'news'. The creation and reticulation of news has long been controlled by large corporations and governments. Pervasive information infrastructure creates the possibility that individuals may be more effectively harnessed as 'spotters' for news organisations. But P2P adds a further dimension: news organisations can be disintermediated. Is there evidence of P2P in the form of Usenet/netnews regathering strength? Is there evidence of informal news networks being used in particular circumstances, such as in times of government suppression of information, and in emergencies arising from natural disasters, warfare and terrorism? Are some forms of P2P, or some features of P2P, particularly suitable for such purposes? Are P2P networks able to bootstrap quickly when conditions arise that call for their use?

8.3 Countermeasures and Their Implications

The threats that copyright-owners and government censors perceive in P2P have resulted in countermeasures. Copyright-owners and their agents are investing in technological protections for digital works. They are participating in P2P networks in order to gather information about them. More directly, music publishers and their agents have been actively polluting the quality of content on P2P networks by introducing similarly-named objects that are incomplete or otherwise spoilt, in an endeavour to reduce the networks' reputation and attractiveness [30]. These measures may or may not be effective in achieving their aims, and may or may not have side-effects.

Research questions include: To what extent are P2P networks vulnerable to pollution attacks? Is copyright-owners' usage of masquerade and pollution techniques reducing the perceived quality and reliability of files shared using P2P networks? If copyright-owners can undermine P2P networks, are they more generally vulnerable to attackers? Are technological protections for copyright objects effective in preventing unauthorised access? Are technological protections for copyright objects harming the exercise of legal rights by consumers, such as licensed use, and use under fair use / fair dealing provisions?

The network equivalent of a field study can be used to investigate many of these questions. Engineering research is then needed in order to gain insights into the extent to which the effects of these attacks can be mitigated. Questions about user perceptions need to be pursued using survey and interpretivist techniques.

Governments are also implementing measures to protect themselves. One example is the use of proxy-servers to block content (including not only comments on the regime in the People's Republic of China, but also pictures of breasts and genitalia in libraries in the U.S.A. and schools in Australia). Another is attempts to identify and track devices and users, and to facilitate their identification and tracking. Research questions include: Are proxy-server techniques effective for censorship, and to what extent are they being used? Do proxy-server techniques have significant side-effects? To what extent are devices identifiable? To what extent are individual users identifiable? How easy is it to circumvent techniques to facilitate identification and tracking? In particular, are proxy-server techniques effective in enabling individuals to achieve anonymisation, and to what extent are they being used?

Governments already have considerable powers available to enable them to counter-attack against impacts of P2P networks that they perceive to be negative. Copyright-owners have not been in as strong a position, because until very recently copyright breach was a purely civil matter, and discovery processes were limited. Lobbying by powerful associations of copyright-owners, however, has resulted in dramatic change over the last decade. Some forms of

copyright breach have been criminalised; and new powers have been granted, such as the U.S. Digital Millennium Copyright Act (DMCA) provisions and the copycat provisions in some other countries (e.g. Lunney 2001), and Anton Piller orders that have recently arisen as an extension to the common law [48]. These are resulting in substantial impositions on the operations and costs of ISPs, on the availability of data, and on consumers. They have been supplemented by aggressive communications by lawyers on behalf of copyright-owners (using so-called 'nastygrams'), which frequently adopt the pretence that they have far more powers than is actually the case.

Research questions include: To what extent are new powers being used by corporations whose interests are negatively affected by the use of P2P networks? To what extent is the use of those powers effective? To what extent do those powers have negative side-effects?

Many business disciplines, including information systems, tend to avoid policy studies, ceding the domain to the applied social sciences and humanities. A variety of research techniques can be applied. Given that P2P is still new, there is scope for application of non-empirical techniques such as conceptual research, simulation modelling, scenario-building and game- or role-playing. Some of the questions are susceptible to survey techniques, and insights can be gained into many of them using interpretivist methods. Field and laboratory experimentation and engineering techniques have much to offer in establishing the extent to which surmised impacts and implications are technically feasible, and empirically evident.

8.4 Legal Aspects

Parliaments have the responsibility to enact and refine legislation which will establish and maintain balances among competing interests. P2P has created an array of challenges that need to be addressed. Preliminary surveys are in [5] and [26].

Questions in this area include: In what ways is current law unsatisfactory in the new context created by P2P technologies? Are some forms of regulation necessary in order to address the perceived negative aspects of P2P, and the negative impacts of countermeasures against them? In what ways might the law be adapted in order to overcome the problems both in the law and in society and the economy? Does P2P increase the need for some form of supra-jurisdictional law that transcends national boundaries as the Internet and P2P traffic do?

9 Conclusions

It is possible that P2P architectures may be in the process of replacing client-server in much the same way that client-server earlier replaced master-slave / star network topologies. If so, it is vital that the management disciplines quickly get to grips with the emergent order of things. Alternatively, the superiority of P2P architectures may be limited to particular contexts. But even in this case, it is clear that P2P's niches include some applications of considerable importance to eBusiness. The reticulation of audio and vide files for entertainment, information and education has attracted most of the media attention, but other examples include the rapid dissemination of software fixes and of data about new items of malware and how to detect and counter them.

Computer science research is delivering deep technical understanding of P2P technology. To date progress in that field has not been matched by equivalent depth in the appreciation of P2P's applications, business significance, impacts and implications. These aspects of P2P require the commitment of research resources within information systems and other management disciplines.

Given the wide array of other topics clamouring for attention from eBusiness researchers, it is important that the overall research program exhibit a reasonable degree of coherence, and that individual projects be conducted in a manner that is effective and reasonably efficient. The research needs to be applied and instrumentalist in orientation, and reports of research outcomes needs to be accessible. Developers, copyright-dependent corporations, government agencies, regulators and consumer advocates all need information about the nature and operation of P2P, in order to develop coherent strategies and action plans that are suitable to the new context.

The purpose of this paper has been to survey this vibrant and rapidly-changing domain, in order to identify research opportunities and provide them with context. A research agenda has been constructed, comprising the dimensions of domain segmentation and research techniques. The limited results reported to date indicate that there is scope for a great deal of valuable research to be undertaken, across all segments of P2P, from technical aspects to organisational, strategic, social, economic and legal studies. Many research techniques are applicable, and a tentative mapping has been offered between research questions and research techniques.

It is contended that this preliminary work on an agenda for P2P research in the eBusiness arena imposes some order on the chaos, and hence provides a basis for a degree of coherence among research activities that are inevitably highly dispersed both geographically and across disciplines. The agenda also contributes to effectiveness and efficiency by suggesting appropriate techniques to apply to the various categories of research question, and by enabling improved discoverability of research outcomes.

The research agenda presented in this paper needs to be subjected to constructive criticism, and refined and extended. The growing body of research reports needs to be identified, catalogued in accordance with this or some improved framework, exploited, and extended. Gaps in research activities need to be identified, and programs and projects developed to address them. These activities are vital if the eBusiness disciplines are to fulfil their potential in the important research domain of peer-to-peer architectures.

P2P Web-Sites

- BigChampagne, at http://www.bigchampagne.com/
- BitTorrent, at http://bittorrent.com/
- Cachelogic, at http://www.cachelogic.com/
- DES cracker, at http://www.eff.org/Privacy/Crypto/Crypto_misc/DESCracker/
- eDonkey, at http://www.edonkey2000.com/
- Edutella, at http://edutella.jxta.org/
- eMule, at http://en.wikipedia.org/wiki/EMule
- FastTrack, at http://en.wikipedia.org/wiki/FastTrack
- fightaids@home, at http://fightaidsathome.scripps.edu/
- folding@home, at http://folding.stanford.edu/
- Freenet, at http://en.wikipedia.org/wiki/Freenet
- genome@home, at http://www.stanford.edu/group/pandegroup/genome/
- Gnutella, at http://en.wikipedia.org/wiki/Gnutella
- Grokster, at http://en.wikipedia.org/wiki/Grokster
- Groove, at http://www.groove.net/
- iMesh, at http://www.imesh.org/
- Internet2 Peer-to-Peer Working Group, at http://p2p.internet2.edu/apps_list.html
- Kademlia, at http://en.wikipedia.org/wiki/Kademlia
- Kazaa Media Desktop (KMD), at http://en.wikipedia.org/wiki/Kazaa
- Lionshare, at http://lionshare.its.psu.edu/main/
- MLDonkey, at http://en.wikipedia.org/wiki/MLDonkey
- Morpheus, at http://en.wikipedia.org/wiki/Morpheus_%28computer_program%29
- Napster, at http://en.wikipedia.org/wiki/Napster
- O'Reilly OpenP2P (Distributed Computation), at http://www.openp2p.com/pub/t/73
- O'Reilly OpenP2P (Distributed Search Engines), at http://www.openp2p.com/pub/t/74
- O'Reilly OpenP2P (File-Sharing), at http://www.openp2p.com/pub/t/75
- O'Reilly OpenP2P (Messaging Frameworks), at http://www.openp2p.com/pub/t/78
- Overnet, at http://en.wikipedia.org/wiki/Overnet
- PeerEnabler, at http://www.joltid.com/index.php/peerenabler/
- SETI@home, at http://setiathome.ssl.berkeley.edu/project.html
- Skype, at http://www.skype.com/
- Slyck, at http://www.slyck.com/
- TeleGeography, at http://www.telegeography.com/

References

[1] D.P. Anderson, J. Cobb, E. Korpela, M. Lebofsky and D. Werthimer. (2002, November). SETI@home: An Experiment in Public-Resource Computing. Communications of the ACM. [Online]. Vol. 45, No. 11 pp. 56-61. Available: http://setiathome.ssl.berkeley.edu/cacm/cacm.html.

[2] R. Anderson. (1997, June). The Eternity Service Cambridge University Computer Laboratory. [Online]. Available: http://www.cl.cam.ac.uk/users/rja14/eternity/eternity.html.

[3] S. Androutsellis-Theotokis and D. Spinellis (2004, December). A Survey of Peer-to-Peer Content Distribution Technologies. ACM Computing Surveys. [Online]. Vol. 36, No. 4, pp. 335-371. Available: http://www.spinellis.gr/pubs/jrnl/2004-ACMCS-p2p/html/AS04.html.

[4] A. Berson, Client/Server Architecture. McGraw-Hill, 2nd edition, 1996

[5] N. Blackmore (2004, March). Peer-To-Peer Networks: The Legal and Technological Challenges for Copyright Owners. N.S.W. Society for Computers and the Law. [Online]. No. 55. Available: http://www.nswscl.org.au/journal/55/Blackmore.html.

[6] Cachelogic. (2005). Trends and Statistics in Peer-to-Peer. Cachelogic Inc.. [Online]. Available: http://www.cachelogic.com/research/p2p2005.php.

[7] D.L. Chaum, Untraceable electronic mail, return addresses, and digital pseudonyms, Communications of the ACM, Vol. 24, No. 2, pp. 84-88, 1982.

[8] R.J. Chevance, Server Architectures : Multiprocessors, Clusters, Parallel Systems, Web Servers, Storage Solutions, Digital Press, 2004.

[9] J. Chuang. (2004). Economics of Peer-to-Peer Systems. Summer Institute on Peer-to-Peer Computing, Academia Sinica, [Online]. Available: http://p2pecon.berkeley.edu/ppt/p2pecon-sinica.pdf.

[10] I. Clarke, O. Sandberg, B. Wiley and T.W. Hong, Freenet: A Distributed Anonymous Information Storage and Retrieval System in Designing Privacy Enhancing Technologies: International Workshop on Design Issues in Anonymity and Unobservability, (H. Federrath Ed.). Lecture Notes in Computer Science, 2000.
[11] I. Clarke, T.W. Hong, S.G. Miller and O. Sandberg (2002, January-February). Protecting Free Expression Online with Freenet. IEEE Internet Computing. [Online]. Vol. 6, No. 1 pp. 40-49. Available: http://www.doc.ic.ac.uk/~twh1/academic/papers/ieee-final.pdf.
[12] R. Clarke. (2003, December). File-Discovery and File-Sharing Technologies (aka Peer-to-Peer or P2P): MP3, Napster and Friends, and Their Impact on E-Publishing of Music, and Other Digital Objects. Xamax Consultancy Pty Ltd. [Online]. Available: http://www.anu.edu.au/people/Roger.Clarke/EC/FDST.html.
[13] R. Clarke. (2004, June). Open Source Software and Open Content As Models for eBusiness. Proceedings of the 17th International eCommerce Conference, Bled, Slovenia. [Online]. Available: http://www.anu.edu.au/people/Roger.Clarke/EC/Bled04.html.
[14] R. Clarke. (2004, November). Peer-to-Peer (P2P) - An Overview, Working Paper. Xamax Consultancy Pty Ltd. [Online]. Available: http://www.anu.edu.au/people/Roger.Clarke/EC/P2POview.html.
[15] B. Cohen. (2003, May). Incentives Build Robustness in BitTorrent, Working Paper. BitTorrent.com. [Online]. Available: http://bittorrent.com/bittorrentecon.pdf.
[16] C. Dubnicki, C. Ungureanu and W. Kilian. (2004, June). FPN: A Distributed Hash Table for Commercial Applications. Proceedings of the Conference HPDC-13,Honolulu, Hawaii USA. [Online]. Available: http://hpdc13.cs.ucsb.edu/papers/184.pdf.
[17] EFF, Cracking DES, Secrets of Encryption Research, Wiretap Politics and Chip Design, Electronic Frontiers Foundation, O'Reilly & Associates, 1998.
[18]E. Felten. (2004, December). TinyP2P: The World's Smallest P2P Application. freedom-to-tinker.com. [Online]. Available: http://www.freedom-to-tinker.com/tinyp2p.html.
[19] W. Felter. (2002, January). Design Choices in P2P Infrastructure. IBM Austin Research Laboratory, slide-set. [Online]. Available: http://www.internet2.edu/presentations/20020130-P2P-Felter.htm.
[20] Gong. (2005, July). Identifying P2P users using traffic analysis. Security Focus Inc. [Online]. Available: http://www.securityfocus.com/print/infocus/1843.
[21] A. Habib and J. Chuang. (2004, June). Incentive Mechanism for Peer-to-Peer Media Streaming. Proceedings of the 12th IEEE International Workshop on Quality of Service (IWQoS'04). [Online]. Available: http://p2pecon.berkeley.edu/pub/HC-IWQOS04.pdf.
[22] J., Howe. (2003, October). BigChampagne is Watching You. Wired. [Online]. No. 11.10. Available: http://www.wired.com/wired/archive/11.10/fileshare.html.
[23] T. Karagiannis, A. Broido, N. Brownlee, K.C. Claffy and M. Faloutsos. (2004, November-December). Is P2P dying or just hiding?. Proceedings of Globecom. [Online]. Available: http://www.cs.ucr.edu/~tkarag/papers/gi04.pdf.
[24] R.S.R. Ku. (2005, February). Grokking Grokster. Case Legal Studies, Research Paper No. 05-5. [Online]. Available: http://ssrn.com/abstract=675856.
[25] J.F. Kurose and K.W. Ross, Computer Networking: A Top-Down Approach Featuring the Internet, Pearson Education, 2004.
[26] E. Lawrence, J. Lawrence and G. Culjak. (2006, April). Legal and Technical Issues Management Framework for Peer-to-Peer Networks. Journal of Theoretical and Applied Electronic Commerce Research. [Online]. Vol. 1, No. 1, pp. 32-41. Available: http://www.jtaer.com.
[27] N. Leibowitz, M. Ripeanu and A. Wierzbicki, Deconstructing the Kazaa Network. 3rd IEEE Workshop on Internet Applications (WIAPP'03), Santa Clara, CA, 2003.
[28] J. Liang, R. Kumar and K.W. Ross. (2004, May), Understanding KaZaA. Working Paper, Polytechnic University, New York. [Online]. Available: http://cis.poly.edu/~ross/papers/UnderstandingKaZaA.pdf.
[29] J. Liang, R. Kumar and K.W. Ross. (2004, September). The KaZaA Overlay: A Measurement Study. Working Paper, Polytechnic University, New York. [Online]. Available: http://cis.poly.edu/~ross/papers/KazaaOverlay.pdf.
[30] J. Liang, R. Kumar, Y. Xi and K.W. Ross. (2005). Pollution in P2P file sharing systems. Proceedings of IEEE INFOCOM. [Online]. Available: http://cis.poly.edu/~ross/papers/pollution.pdf.
[31] B. Loban. (2004, October). Between rhizomes and trees: P2P information systems. First Monday. [Online]. Vol. 9, No. 10. Available: http://www.firstmonday.org/issues/issue9_10/loban/index.html.
[32] G.S. Lunney Jr., The Death of Copyright: Digital Technology, Private Copying, and the DMCA**,** Virginia Law Review, Vol. 87, September, 2001.
[33] P. Lyman and H.R. Varian. (2003). How Much Information?. School of Information Management and Systems, University of California at Berkeley. [Online]. Available: http://www.sims.berkeley.edu/research/projects/how-much-info-2003/.
[34] S. Mann. (1997, February). Wearable Computing: A First Step Toward Personal Imaging. Computer. [Online]. Vol. 30, No. 2. Available: http://www.wearcam.org/ieeecomputer/r2025.htm.
[35] N. Minar and M. Hedlund. (2001). A Network of Peers – Peer-to-Peer Models Through the History of the Internet, Chapter 1 of [38]. [Online]. Available: http://www.oreilly.com/catalog/peertopeer/chapter/ch01.html.
[36] K. Nakauchi, Y. Ishikawa, H. Morikawa and T. Aowama, Exploiting Semantics in Unstructured Peer-to-Peer Networks, IEICE Transactions in Communications, Vol. E87-B, pp. 1806-1817, July, 2004.
[37] OECD. (2004, July). Peer to Peer Networks in OECD Countries. OECD, Paris. [Online]. Available: http://www.oecd.org/dataoecd/55/57/32927686.pdf.
[38] A. Oram (Ed.). (2001). Peer-to-Peer: Harnessing the Power of Disruptive Technologies. O'Reilly. [Online]. Available: http://www.oreilly.com/catalog/peertopeer/index.html.

[39] J.A. Pouwelse, P. Garbacki, D.H.J. Epema and H.J. Sips. (2005, February). The Bittorrent P2P File-sharing System: Measurements and Analysis. Proceedings of the 4th International Workshop on Peer-to-Peer Systems (IPTPS'05). [Online]. Available: http://www.isa.its.tudelft.nl/~pouwelse/Bittorrent_Measurements_6pages.pdf.
[40] A. Preston. (2002). Peer-to-peer: an overview of a disruptive technology. Internet2 Peer-toPeer Working Group, slide-set. [Online]. Available: http://www.terena.nl/conferences/tnc2002/Slides/sl8b1.ppt.
[41] RFC1. (1969, April). Host Software. RFC1, [Online]. Available:http://www.faqs.org/rfcs/rfc1.html.
[42] J. Risson and T. Moors, Survey of Research towards Robust Peer-to-Peer Networks: Search Methods, University of N.S.W., Sydney, Technical Report UNSW-EE-P2P-I-I, September 2004.
[43] K.W. Ross. (2004). Recommended Reading in P2P Networking Theory. Polytechnic University, New York. [Online]. Available: http://cis.poly.edu/~ross/p2pTheory/P2Preading.htm.
[44] M. Roussopoulos, M. Baker, D.S.H. Rosenthal, T.J. Giuli, P. Maniatis and J. Mogul. (2004, February). 2 P2P or Not 2 P2P?. Proceedings of IPTPS 2004. [Online]. Available: http://www.eecs.harvard.edu/~mema/publications/iptps2004.pdf.
[45] SEI. (1997). Client/Server Software Architectures--An Overview. Software Engineering Institute, Carnegie-Mellon University. [Online]. Available: http://www.sei.cmu.edu/str/descriptions/clientserver_body.html.
[46] M. Skala. (2005, January). MoleSter 0.0.4 - now 6 lines, 466 bytes. [Online]. Available: http://ansuz.sooke.bc.ca/software/molester/.
[47] H.A. Smith, J. Clippinger and B. Konsynski, Riding the Wave: Discovering the Value of P2P Technologies, Commun. Assoc. Infor. Syst., Vol. 11, No. 4, 94-107, January, 2003.
[48] N. Suzor. (2004). Privacy v Intellectual Property Litigation: Preliminary Third Party Discovery on the Internet, Australian Bar Review. Issue 25 p. 228. [Online]. Available: http://ssrn.com/abstract=627786
[49] TeleGeography. (2005). Global Internet Geography. Primetrica Inc., [Online]. Available: http://www.telegeography.com/ee/free_resources/reports/gig/gig_exec_sum.php.
[50] E. Truch, J.-N. Ezingeard and D.W. Birchall. (2000). Developing a relevant research agenda in Knowledge Management – bridging the gap between knowing and doing, In Proceedings of the Eighth European Conference on Information Systems (H.R. Hansen, M. Bichler, H. Mahrer, Eds.). [Online]. pp. 694-700, Vienna. Available: http://csrc.lse.ac.uk/asp/aspecis/20000190.pdf.
[51] M. Waldman, A.D. Rubin and L.F. Cranor. (2000, August). Publius: A robust, tamper-evident, censorship-resistant, web publishing system. Proceedings of the 9th USENIX Security Symposium. [Online]. Available: http://cs1.cs.nyu.edu/~waldman/publius/publius.pdf.
[52] Y. Wand and R. Weber, Research Commentary: Information Systems and Conceptual Modelling – A Research Agenda, Information Systems Research, Vol. 13, No. 4, December, 2002.
[53] H. Wen. (2002, September). Internet Radio the P2P Way. O'Reilly P2P.com. [Online]. Available: http://www.openp2p.com/pub/a/p2p/2002/09/24/p2pradio.html.
[54] M. Xie. (2003, September). P2P Based on Distributed Hash Table. Department of Computer Science, University of Ottawa. [Online]. Available: http://www.site.uottawa.ca/~mxie/academic/bak/DHT.pdf.